Joint 45th International Symposium on Robotics (ISR) and 8th German Conference on Robotics (ROBOTIK), Munich, June 2014.

# 3D Local Multiresolution Grid for Aggregating Omnidirectional Laser Measurements on a Micro Aerial Vehicle

David Droeschel and Sven Behnke
Autonomous Intelligent Systems Group, University of Bonn, Germany

## Abstract

Micro aerial vehicles (MAV) pose a challenge in designing sensory systems and algorithms due to their size and weight constraints and limited computing power. We present an efficient 3D multiresolution map that we use to aggregate measurements from our lightweight continuously rotating laser scanner. We estimate robot motion by means of visual odometry and scan registration, aligning consecutive 3D scans with an incrementally built map. Hence, we are able to aggregate scans over time and increase the density of the map.

By using local multiresolution, we gain computational efficiency by having a high resolution in the near vicinity of the robot and a lower resolution with increasing distance from the robot, which correlates with sensor characteristics in relative distance accuracy and measurement density. Compared to uniform grid-based maps, multiresolution leads to the use of fewer grid cells without loosing information and consequently results in lower computational costs. Hence, we are able to efficiently register new 3D scans with the map, utilizing the multiresolution property in a coarse-to-fine approach.

In experiments, we compare the laser-based motion estimate with ground-truth from a motion capture system as well as a state-of-the-art registration method. Overall, our approach allows for building accurate 3D obstacle maps and estimating the MAV trajectory in real-time by 3D scan registration.

## 1 INTRODUCTION

Micro aerial vehicles (MAV) such as quadrotors have attracted attention in the field of aerial robotics. Their size and weight limitations pose a problem in designing sensory systems for these robots, however. Most of today's MAVs are equipped with ultra sound sensors and camera systems due to their minimal size and weight. While these small and lightweight sensors provide valuable information, they suffer from a limited field-of-view and are sensitive to illumination conditions. Only few systems [17, 5, 1, 15] are equipped with 2D laser range finders (LRF) that are used for navigation.

In contrast, we build a continuously rotating laser scanner that is minimalistic in terms of size and weight and thus particularly well suited for obstacle perception and localization on MAVs, allowing for environment perception in all directions.

We use a hybrid multiresolution map that stores occupancy information and the respective distance measurements. Measurements are stored in grid cells with increasing cell size from the robot's center. Thus, we gain computational efficiency by having a high resolution in the close proximity to the sensor and a lower resolution with increasing distance, which correlates with the sensor's characteristics in relative distance accuracy and measurement density.

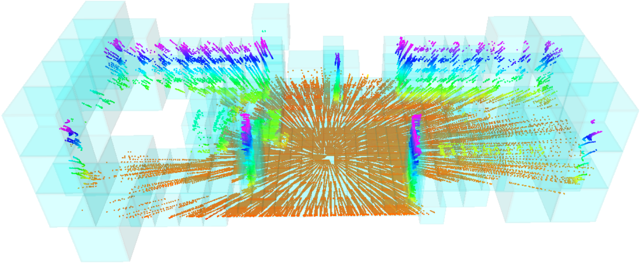
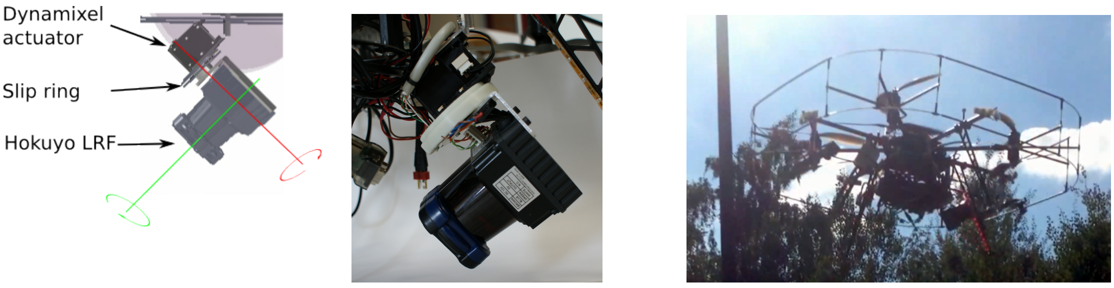


Figure 1: The grid-based multiresolution map with a higher resolution in the close proximity to the sensor and a lower resolution with increasing distance. The point color encodes the distance from the ground.

Compared to uniform grid-based maps, multiresolution leads to the use of fewer grid cells without loosing information and consequently results in lower computational costs. **Fig. 1** shows our multiresolution grid-based map.

Aggregating measurements from consecutive time steps necessitates a robust and reliable estimation of the sensor motion. Thus, we use the point-based map representation to obtain an estimate of the sensor motion between consecutive 3D scans by scan registration. Since laser-based ego-motion estimation relies on structure in the scene, it works best in scenarios where Global Navigation Satellite Systems (GNSS) are not available, like in indoor or urban environments.



(a) (b) (c)

Figure 2: (a) CAD drawing of the continuously rotating laser scanner with the two rotation axes. The Hokuyo 2D LRF is mounted on a bearing and rotated around the red axis. (b) A photo of the assembled laser scanner. (c) The 3D laser scanner mounted on our multicopter.

The remainder of the paper is structured as follows. After a brief discussion of related work, the sensor setup is described. The multiresolution map is introduced in Section 4, followed by the scan registration in Section 5. In the experiment section, we evaluate our approach and compare it to ground-truth data from a motion capture system as well as a state-of-the-art registration method.

## 2 Related Work

For mobile ground robots, 3D laser scanning sensors are widely used due to their accurate distance measurements even in bad lighting conditions and their large field-of-view (FoV). For instance, autonomous cars often perceive obstacles by means of a rotating laser scanner with a 360° horizontal FoV, allowing for detection of obstacles in every direction [4, 9].

Up to now, such 3D laser scanners are rarely used on lightweight MAVs, due to their payload limitations. Instead, two-dimensional laser range finders [6, 17, 5, 1, 15, 7] are used, which restricts the field-of-view to the two-dimensional measurement plane of the sensor.

A similar setup to ours is described by Scherer and Cover et al. [12, 3]. Their MAV is used to autonomously explore rivers using visual localization and laser-based 3D obstacle perception. In contrast to their work, we aggregate consecutive laser scans in our multiresolution map by 3D scan registration.

For mobile ground robots, some approaches have been proposed to estimate the motion of a robot by means of 3D scan registration [11, 8, 14]. Most of these approaches are derived from the Iterative Closest Points (ICP) algorithm [2]. Generalized-ICP [14] unifies the ICP formulation for various error metrics such as point-to-point, point-to-plane, and plane-to-plane.

We aim at perceiving as much of the surroundings as possible in order to obtain almost omnidirectional obstacle detection. Distance measurements are aggregated in a 3D grid-based map at multiple resolutions and acquired 3D scans are registered incrementally with the map.

## 3 Sensor Setup

Our continuously rotating 3D laser scanner consists of a Hokuyo UTM-30LX-EW 2D laser range finder (LRF) which is rotated by a Dynamixel MX-28 servo actuator to gain a three-dimensional FoV. As shown in **Fig. 2**, the scanning plane is parallel to the axis of rotation, but the heading direction of the scanner is twisted slightly away from the direction of the axis—in order to enlarge its field-of view. The 2D LRF is electrically connected by a slip ring, allowing for continuous rotation of the sensor. The sensor is mounted on our multicopter (Fig. 2c) pitched downward by 45° in forward direction, which places the core of the robot upwards behind the sensor. Hence, the sensor can measure in all directions, except for a conical blind spot pointing upwards behind the robot.

The 2D laser scanner has a size of $62 \times 62 \times 87.5\,\mathrm{mm}$ and a weight of 210 g. Together with the actuator (72 g) and the slip ring, the total weight of the 3D scanner is approximately 400 g.

The Hokuyo LRF has an apex angle of 270° and an angular resolution of 0.25°, resulting in 1080 distance measurements per 2D scan, called a *scan line*. The Dynamixel actuator rotates the 2D LRF at a rate of one rotation per second, resulting in 40 scan lines and 43,200 distance measurements per full rotation. Slower rotation is possible if a higher angular resolution is desired. For our setup, a half rotation leads to a full 3D scan of most of the environment (**Fig. 3**). Hence, we can acquire 3D scans with up to 21,600 points with 2 Hz.

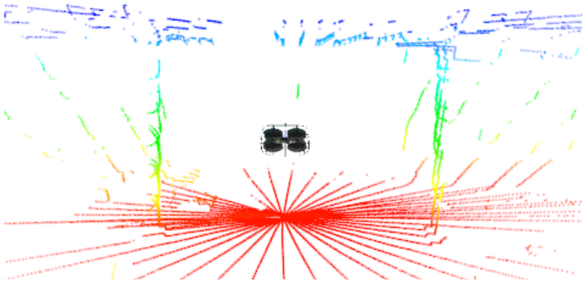


Figure 3: A 3D scan of an indoor environment acquired with our continuously rotating laser scanner. The color of the points encodes the distance from the ground plane.

# 4 Local Multiresolution Map

Distance measurements from the sensor are accumulated in a 3D multiresolution map with increasing cell sizes from the robot center. The representation consists of multiple robot-centered 3D grid-maps with different resolutions. On the finest resolution, we use a cell length of 0.25 m. Each grid-map is embedded in the next level with coarser resolution and doubled cell length.

We use a hybrid representation, storing 3D point measurements along with occupancy information in each cell. Point measurements of consecutive 3D scans are stored in fixed-sized circular buffers, allowing for point-based data processing and facilitating efficient nearest-neighbor queries.

**Fig. 4** shows a one-dimensional schematic illustration of the map organization. We aim for efficient map management for translation and rotation. Therefore, individual grid cells are stored in a circular buffer to allow shifting of elements in constant time. We interlace multiple circular buffers to obtain a map with three dimensions. The length of the circular buffers depends on the resolution and the size of the map. In case of a translation of the MAV, the circular buffers are shifted whenever necessary to maintain the egocentric property of the map. In case of a translation equal or larger than the cell size, the circular buffers for respective dimensions are shifted. For sub-cell-length translations, the translational parts are accumulated and shifted if they exceed the length of a cell.

Since we store 3D points for every cell for point-based processing, individual points are transformed in to the cell local coordinate frame when adding, and back to the map's coordinate frame when accessing. Every cell in the map stores a list of 3D points from the current and previous 3D scans. This list is also implemented by a fixed-sized circular buffer. If the capacity of the circular buffer is exceeded, old measurements are discarded and replaced by new measurements.

Since rotating the map would necessitate to shuffle all cells, our map is oriented independent of the MAV orientation. We maintain the orientation between the map and the MAV and use it to rotate measurements when accessing the map.

Besides the scan registration described in the following section, the map is utilized by our obstacle avoidance control using a predictive potential field method to avoid occupied cells [10].

# 5 Scan Registration

We register consecutive 3D laser range scans to estimate the motion of the robot. After acquiring a full 3D scan (i.e., a half rotation), the scan is deskewed to compensate for the sensor motion and aligned to the map by the Iterative Closest Point (ICP) algorithm.

Correspondences are assigned using the point-based representation in the grid-cells and the ICP algorithm estimates a transformation between the scan and the map, describing the displacement between them. We benefit from the multiresolution property of our map, which allows to align a 3D scan in a coarse-to-fine approach. Hence, we start assigning correspondences and estimating the transformation at the coarsest level. The resulting transformation is used as initialization for the registration on the next finer level and so forth.

## 5.1 Scan Undistortion

Since movement of the sensor during acquisition leads to a distortion of the 3D scan, we incorporate a visual odometry estimate from two pairs of wide-angle stereo cameras [13]. This 6D motion estimate is used to assemble the individual 2D scan lines of each a half rotation to a 3D scan. **Fig. 5** illustrates the effect of scan undistortion.

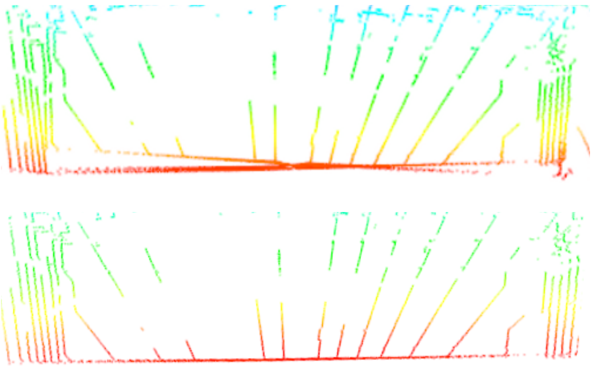


Figure 5: 3D point cloud of an indoor environment with flat ground from a side view. Sensor movement during scan acquisition yields distorted 3D scans (top). We deskew the scan based on the motion estimate (bottom).

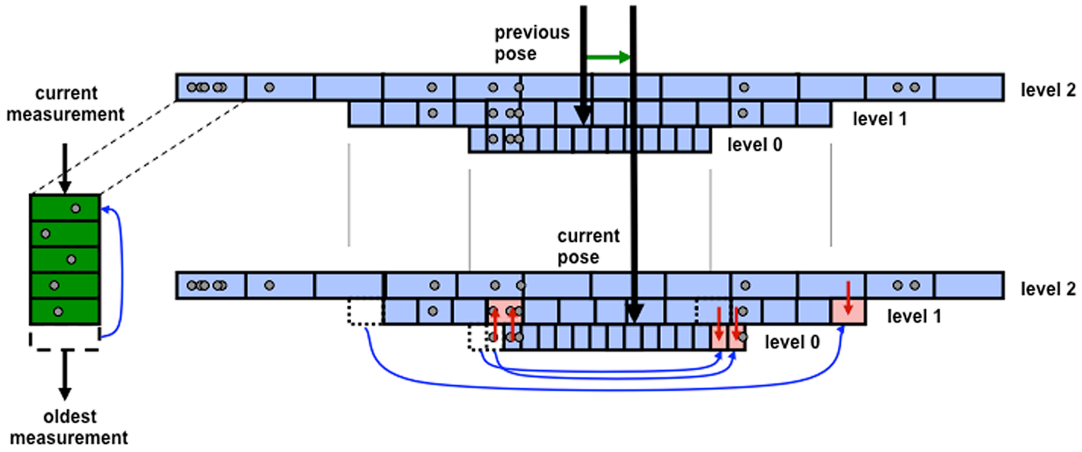


Figure 4: One-dimensional schematic illustration of the hybrid local multiresolution map. Along with the occupancy information, every grid-cell (blue) maintains a circular buffer with its associated measurement points (green). The map is centered around the robot and in case of a robot motion, ring buffers are shifted according to the translational parts of the movement, maintaining the egocentric property of the map. Cells at coarser levels are used to retain points from vanishing cells at finer levels and to initialize newly added cells (red arrows).

## 5.2 Data Association

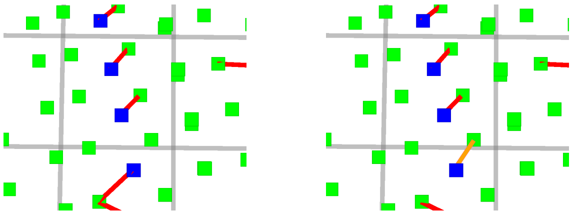

(a) (b)

Figure 6: Assigning point correspondences. (a) For every point of a 3D scan (blue), a corresponding map point (green) is initially assigned from the cell's point list (red line). (b) If the distance to neighboring cells is smaller than the distance to the initial assignment, closer points might be found in the neighboring cell (orange line).

When using the ICP algorithm for scan registration, corresponding points between the model and the current point cloud are assigned, usually by building a space-partitioned data structure from the model point cloud. In contrast, we continuously maintain our data structure for efficient nearest-neighbor queries to assign correspondences. Every point from a newly acquired 3D scan is directly assigned to a cell in the map in constant time. The closest point in terms of the Euclidean distance from the point list of this cell is initially assigned as corresponding point.
As illustrated in **Fig. 6**, points in neighboring cells might be closer to the measured point than the initially assigned point. Consequently, we extend the search to neighboring cells, if the distance to the initial assignment is larger than the distance to the border of a neighboring cell.
Since acquired 3D scans of the scene and the aggregated local map differ in terms of structure and point density, especially when parts of the scene have previously been occluded, individual assigned correspondences can be incorrect. These incorrect correspondences distort the transformation estimation and need to be filtered. Thus, we reject correspondences by the following criteria:

- *Asymmetric correspondences:* We check for symmetry in the assignment, i.e., for a correspondence from a scan point $d_i$ to a map point $m_i$, we check if $d_i$ is the closest point to $m_i$ in the scan point cloud. Otherwise, the correspondence is rejected.

- *One-to-many correspondences:* In case one point in the map corresponds to many points in the scan cloud, we keep the correspondence with the least distance and reject the remaining.

- *Correspondence trimming:* Correspondences are rejected by only considering $\theta_t$ percent of the assigned correspondences with the least distance.

- *Distance rejection:* Correspondences are rejected if farther away than $\theta_d$.

## 5.3 Transformation Estimation

With $N$ assigned corresponding point pairs $(m_i, d_i)$, we determine the displacement between the points of a scan $d_i$ and the map points $m_i$ by finding a rigid transformation $T$ that minimizes

Table 1: ATE and run-time of our registration method, in comparison to visual odometry (VO), and GICP.

| | ATE (m) | | | | | run-time (ms) | | |
|---|---|---|---|---|---|---|---|---|
| | RMSE | mean | median | std | max | mean | std | max |
| VO | 0.151 | 0.134 | 0.129 | 0.059 | 0.324 | | | |
| GICP | 0.033 | 0.030 | 0.030 | 0.013 | **0.079** | 1432.96 | 865.24 | 5673 |
| ours | **0.030** | **0.028** | **0.026** | 0.015 | 0.093 | **311.28** | 90.31 | **376** |

$$E(\mathbf{T}) = \sum_{i=1}^{N} ||\mathbf{m}_i - \mathbf{T}d_i||^2 \tag{1}$$

using a closed-form singular value decomposition (SVD) algorithm [2].

In each ICP iteration, correspondences are reassigned, the transformation that aligns these best is applied to the scan, and the following termination criteria are checked:

- $E(\mathbf{T})$ is smaller than a given threshold $\theta_r$,
- the difference between $T_k$ and $T_{k-1}$ is smaller than $\theta_\epsilon$, or
- the number of iterations exceeds $\theta_i$,

where $T_k$ and and $T_{k-1}$ are the estimated transformations from the current and the previous iteration, respectively.

## 6 Experiments

In a first experiment, we evaluate the accuracy of the scan registration in an indoor motion capture (MoCap) system. It provides accurate pose information of the MAV at high frame rates (100 Hz) but is restricted to a small capture volume of approximately 2×2×3 m.

As error metric, the absolute trajectory error (ATE) is computed, based on the estimated and the ground-truth trajectory from the MoCap system. The reference implementation provided by Sturm et al. [16] was used.

Throughout the experiments, five levels are used for the multiresolution map with a cell length of 0.125 m at the finest level, yielding a cell length of 2 m at the coarsest level.

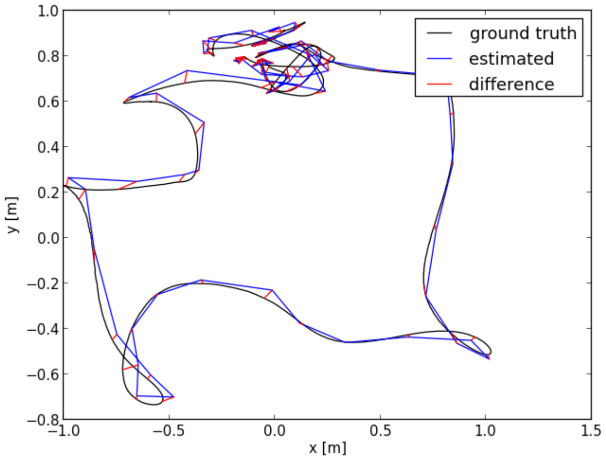


Figure 7: Absolute trajectory error of the scan registration using the multiresolution map (blue) compared to ground-truth data from the MoCap system (black). Points of the trajectory are projected on the xy-plane.

The parameters $\theta_r$, $\theta_i$ and $\theta_\epsilon$ are manually determined. In this experiment, $\theta_r$=1 cm, $\theta_\epsilon$=0.001 cm, $\theta_i$=15, $\theta_d$=1 m, and $\theta_t$=80% showed best results.

The data set for evaluation is a 50 seconds flight sequence containing 100 3D scans, where the MAV was controlled by a human operator, taking off and landing at two locations in the MoCap volume. **Fig. 7** shows the ATE of our multiresolution scan registration method, comparing it to the trajectory of the MoCap system.

In quantitative experiments, we compare our method to a state-of-the-art registration method, the Generalized-ICP [14]. In addition, we evaluate the accuracy of the visual odometry alone that is used to undistort acquired 3D scans. The mean, standard deviation and maximum ATE of all three methods are summarized in **Table 1**. The results indicate that both scan registration methods improve the motion estimate from the visual odometry and that the trajectory generated by our method has a slightly lower ATE compared to the Generalized-ICP. The run-times of both methods for this experiment are also summarized in Table 1, showing that Generalized-ICP is computationally much more expensive than our method.

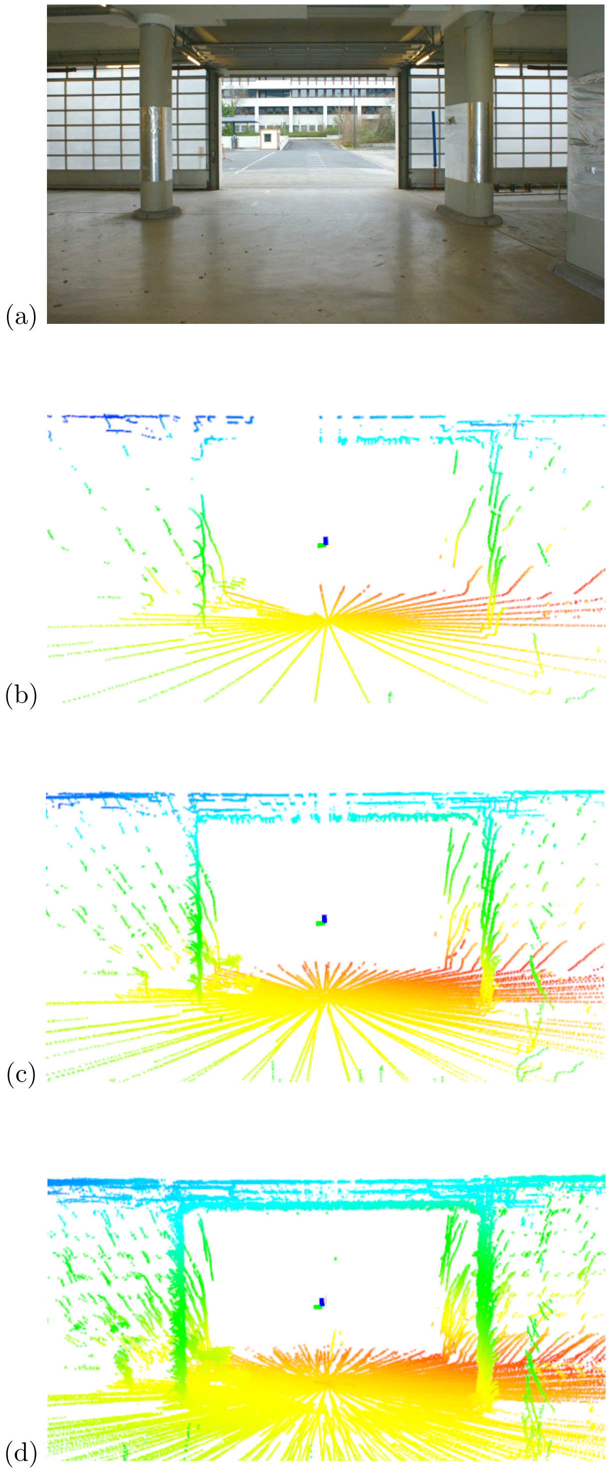


Figure 8: Photo of the scene of the second experiment (a). Aggregated 3D map after 1 scan (b), 5 scans (c), and 10 scans (d) using the multiresolution scan registration.

In a second experiment, we acquire data in a parking garage. **Fig. 8**(a) shows a photo of the environment. Fig. 8 shows the point-based representation of the map at different time steps, accounting for an increasing density of the map after adding consecutive 3D scans. Since ground-truth data, e.g., from a MoCap system was not available in this experiment, we evaluate the different methods by inspecting the variation of points in a planar area. **Fig. 9** shows a part of the floor in the resulting point-based representation from a side-view. It can be seen that using scan registration decreases the thickness of the floor significantly, indicating an improved motion estimate. Similar to the results of the first experiment, the resulting floor patch generated by aggregating scans using our method is slightly thinner. Note that in normal operation, scans are only added if they properly align with the map, i.e., $E(\mathbf{T})$ is smaller than $\theta_r$ in Eq. (1). For this experiment, we added every scan to the map to have a fair comparison to the motion estimate from the visual odometry.

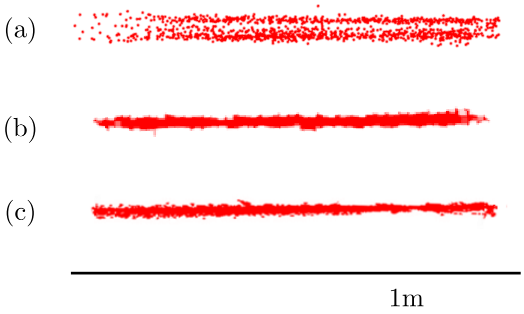


Figure 9: A cut-out part of the floor from a side-view after scan aggregation. (a) using only visual odometry; (b) visual odometry combined with GICP scan registration; (c) visual odometry with our multiresolution scan registration.

## 7 CONCLUSIONS

We presented an efficient 3D multiresolution map that we use for obstacle avoidance and for estimating the motion of the robot. We aggregate measurements from a continuously rotating laser scanner that is particularly well suited for MAVs due to its size and weight.

By using local multiresolution, we gain computational efficiency by having a high resolution in the near vicinity of the robot and a lower resolution with increasing distance from the robot, which correlates with the sensor's characteristics in relative distance accuracy and measurement density.

Scan registration is used to estimate the motion of the robot by aligning consecutive 3D scans to the map in a coarse-to-fine strategy. Hence, we are able to efficiently align new 3D scans with the map and aggregate distance measurements from consecutive 3D scans to increase the density of the map.

In experiments, we compare the laser-based motion estimate with ground-truth from a motion capture system and with the Generalized-ICP, a state-of-the-art registration algorithm. Overall, we can build an accurate 3D obstacle map and can estimate the vehicle's trajectory by 3D scan registration. Compared to the Generalized-ICP, our approach results in a smaller absolute trajectory error and is computational more efficient, allowing to register scans in real-time.

# ACKNOWLEDGMENT

This work has been supported partially by grant BE 2556/7-1 of German Research Foundation (DFG).

# References

[1] A. Bachrach, R. He, and N. Roy. Autonomous flight in unstructured and unknown indoor environments. In *European Micro Aerial Vehicle Conf (EMAV)*, pages 1–8, 2009.

[2] P. J. Besl and N. D. McKay. A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 14(2):239–256, 1992.

[3] H. Cover, S. Choudhury, S. Scherer, and S. Singh. Sparse tangential network (SPARTAN): Motion planning for micro aerial vehicles. In *Robotics and Automation (ICRA), IEEE International Conference on*, 2013.

[4] C. Urmson et. al. Autonomous driving in urban environments: Boss and the urban challenge. *Journal of Field Robotics Special Issue on the 2007 DARPA Urban Challenge, Part I*, 25(8):425–466, June 2008.

[5] S. Grzonka, G. Grisetti, and W. Burgard. Towards a navigation system for autonomous indoor flying. In *Robotics and Automation (ICRA), IEEE International Conference on*, 2009.

[6] S. Grzonka, G. Grisetti, and W. Burgard. A fully autonomous indoor quadrotor. *IEEE Trans. on Robotics*, 28(1):90–100, 2012.

[7] S. Huh, D.H. Shim, and J. Kim. Integrated navigation system using camera and gimbaled laser scanner for indoor and outdoor autonomous flight of uavs. In *Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on*, pages 3158–3163, 2013.

[8] M. Magnusson, T. Duckett, and A. J. Lilienthal. Scan registration for autonomous mining vehicles using 3D-NDT. *Journal of Field Robotics*, 24(10):803–827, 2007.

[9] M. Montemerlo, J. Becker, S. Bhat, H. Dahlkamp, D. Dolgov, S. Ettinger, D. Haehnel, T. Hilden, G. Hoffmann, B. Huhnke, D. Johnston, S. Klumpp, D. Langer, A. Levandowski, J. Levinson, J. Marcil, D. Orenstein, J. Paefgen, I. Penny, A. Petrovskaya, M. Pflueger, G. Stanek, D. Stavens, A. Vogt, and S. Thrun. Junior: The stanford entry in the urban challenge. *Journal of Field Robotics*, 25(9):569–597, 2008.

[10] M. Nieuwenhuisen and S. Behnke. Hierarchical planning with 3d local multiresolution obstacle avoidance for micro aerial vehicles. In *Proceedings of the Joint Int. Symposium on Robotics (ISR) and the German Conference on Robotics (ROBOTIK)*, 2014.

[11] A. Nuechter, K. Lingemann, J. Hertzberg, and H. Surmann. 6D SLAM with approximate data association. In *Robotics and Automation (ICRA), IEEE International Conference on*, pages 242–249, 2005.

[12] S. Scherer, J. Rehder, S. Achar, H. Cover, A. Chambers, S. Nuske, and S. Singh. River mapping from a flying robot: state estimation, river detection, and obstacle mapping. *Autonomous Robots*, 32(5):1–26, May 2012.

[13] J. Schneider, T. Läbe, and W. Förstner. Incremental real-time bundle adjustment for multi-camera systems with points at infinity. In *ISPRS Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume XL-1/W2, 2013.

[14] A. Segal, D. Haehnel, and S. Thrun. Generalized-ICP. In *Proc. of Robotics: Science and Systems (RSS)*, 2009.

[15] S. Shen, N. Michael, and V. Kumar. Autonomous multi-floor indoor navigation with a computationally constrained micro aerial vehicle. In *Robotics and Automation (ICRA), IEEE International Conference on*, pages 2968 –2969, 2011.

[16] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers. A benchmark for the evaluation of rgb-d slam systems. In *Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on*, Oct. 2012.

[17] T. Tomić, K. Schmid, P. Lutz, A. Domel, M. Kassecker, E. Mair, I.L. Grixa, F. Ruess, M. Suppa, and D. Burschka. Toward a fully autonomous UAV: Research platform for indoor and outdoor urban search and rescue. *Robotics Automation Magazine, IEEE*, 19(3):46–56, 2012.